

1 **Both DNA Polymerases δ and ϵ Contact Active and Stalled Replication Forks differently**

2 Chuanhe Yu^{1,2}, Haiyun Gan^{2,3}, Zhiguo Zhang³

3 ¹Department of Biochemistry and Molecular Biology, Mayo Clinic, Rochester, MN, 55905, USA

4 ³Institute for Cancer Genetics, Department of Pediatrics, Department of Genetics and
5 Development, Columbia University, 1130 St. Nicholas Avenue, Irving Cancer Research Center
6 ICRC 407A, New York, NY10032

7

8 ²Equal contribution

9

10

11

12 *Corresponding author: zz2401@cumc.columbia.edu

13 (p): 212-851-4239

14

15 Running title: DNA polymerases configuration on replication fork

16 Key words: DNA replication, replication stress, DNA polymerase, CHIP-ssSeq, strand specific
17 sequencing

18

19

20

21 **Abstract**

22 Three DNA polymerases (Pol α , Pol δ , and Pol ϵ) are responsible for eukaryotic genome
23 duplication. When DNA replication stress is encountered, DNA synthesis stalls until the stress is
24 ameliorated. However, it is not known whether there is a difference in the association of each
25 polymerase with active and stalled replication forks. Here, we show that each DNA polymerase
26 has distinct patterns of association with active and stalled replication forks. Pol α is enriched at
27 extending Okazaki fragments of active and stalled forks. In contrast, although Pol δ contacts the
28 nascent lagging strands of active and stalled forks, it binds to only the matured (and not
29 elongating) Okazaki fragments of stalled forks. Pol ϵ has a greater contact with the nascent
30 ssDNA of leading strand on active forks compared with stalled forks. We propose that the
31 configuration of DNA polymerases at stalled forks facilitate resumption of DNA synthesis after
32 stress removal.

33

34

35

36

37

38

39

40

41 **Introduction**

42 During eukaryotic genome duplication, replication stress is known to cause DNA synthesis to
43 stall until the stress is alleviated. Replication stress includes lesions induced by endogenous and
44 exogenous DNA-damaging agents, ribonucleotide mis-incorporation, and formation of secondary
45 structures or DNA-RNA hybrids (1, 2). To better understand how genome integrity is maintained
46 throughout replication stress, it is critical to determine how replisomes associate with DNA in
47 active and stalled replication forks.

48 In budding yeast, DNA replication initiates at multiple sites, termed autonomously
49 replicating sequences (ARSs) or replication origins. These origins are regulated temporally, with
50 some origins firing early and others firing late in S phase of the cell cycle (3). In budding yeast,
51 one of the primary responses to DNA replication stress is activation of the Mec1 and Rad53
52 kinase signaling cascade, a process equivalent to ataxia telangiectasia mutated– and Rad3-related
53 (ATR) activation in human cells. Activated checkpoint kinases inhibit firing of late replication
54 origins, maintain the stability of stalled replication forks, and help restart DNA synthesis.

55 DNA polymerases α , ϵ , and δ (Pol α , Pol ϵ , and Pol δ) are the main replicative DNA
56 polymerases for eukaryotic nuclear genome, but other proteins are also involved in the process of
57 DNA replication, including the Origin Recognition Complex (ORC) and replicative helicase
58 minichromosome maintenance (MCM) proteins. In G1/S transition, MCM is activated through
59 formation of the CMG complex (Cdc45, Mcm2-7, and GINS complex) and phosphorylation two
60 kinases, CDK and DDK(4). Activated CMG helicase unwinds double-stranded DNA (dsDNA) at
61 origins and recruits the single-stranded DNA (ssDNA)–binding protein RPA; RPA facilitates
62 recruitment of Pol α , which synthesizes RNA primers followed by short DNA chain to initiate

63 leading strands and Okazaki fragment synthesis. Pol ϵ and Pol δ extend Pol α 's products. In
64 budding yeast, based on mutation bias introduced by Pol ϵ and δ mutants, it was proposed that
65 Pol ϵ and δ are responsible for the synthesis of the leading and lagging strands, respectively (5-7).
66 Recently, by mapping ribonucleotides introduced by DNA Pol ϵ and Pol δ mutants genome wide
67 in both budding yeast and fission yeast (8-11), it is deduced by that Pol ϵ and Pol δ are involved in
68 synthesis of leading and lagging strand, respectively. We have shown that Pol ϵ and Pol δ are
69 enriched at nascent leading and lagging strands, respectively (12). However, a recent study of
70 budding yeast suggest that Pol δ is involved in the synthesis of both leading and lagging strands
71 and Pol ϵ is involved in DNA repair (13).

72

73 We and others have shown that proliferating cell nuclear antigen (PCNA), a processivity
74 factor for Pol ϵ and Pol δ , is unloaded from lagging strands of stalled DNA replication forks and
75 that this unloading is regulated by checkpoint kinases (12). Therefore, while replication proteins
76 still associate with replisomes under stress so that DNA synthesis can resume once DNA
77 replication stress is terminated, their contacts with DNA at stalled fork may differ from those
78 with active forks. To analyze the interaction of DNA polymerases with replication forks in
79 budding yeast, we evaluated the association of Pol α , Pol ϵ , and Pol δ with DNA in active and
80 hydroxyurea (HU)-stalled replication forks by using chromatin immunoprecipitation plus strand-
81 specific next-generation DNA sequencing (ChIP-ssSeq). Here, we report an in-depth analysis of
82 protein ChIP-ssSeq datasets, which reveals distinct pattern of interaction of Pol δ and Pol ϵ with
83 DNA at active and stalled replication forks. We suggest that these changes in contact with DNA

84 directly or indirectly help maintain stability of stalled forks and facilitate the resumption of DNA
85 synthesis after amelioration of replication stress.

86

87 **Methods**

88 **Yeast strains**

89 Yeast strains used in this study were derived from W303 (*leu2-3, 112 ura3-1 his3-11, trp1-1,*
90 *ade2-1 can1-100*). Genotypes are listed in Supplemental Table 1.

91

92 **ChIP-ssSeq Procedure**

93 ChIP-ssSeq experiments were performed as described previously (12). Briefly, α factor was used
94 to synchronize yeast cells at G1 (5 μ g/mL and 50 ng/mL for wild-type *BAR* and *bar1* mutant
95 strains, respectively). To analyze the association of proteins with active forks, G1-arrested cells
96 were released into chilled (16°C) YPD medium containing 400 mg/L bromodeoxyuridine (BrdU),
97 and samples were collected at different time points. We treated cells with HU, an inhibitor of
98 ribonucleotide reductase, to deplete cells of dNTP. Thus, HU stalls the progression of DNA
99 replication forks and inhibits the firing of late replication origins (14, 15). To analyze protein
100 association with HU-stalled forks, cells were released into fresh medium containing 400 mg/L
101 BrdU and 0.2M HU for 45 minutes. To perform ChIP, samples were incubated with 1%
102 paraformaldehyde at 25°C for 20 minutes and then quenched with 0.125 M glycine for 5 minutes.
103 Cells were lysed with glass beads, and the chromatin pellet was washed and sonicated to shear

104 DNA to an average fragment size of about 200-400 bp. Sheared chromatin was
105 immunoprecipitated with anti-FLAG antibody F1804 or anti-RPA antibody (gift of Dr Steven
106 Brill). After extensive washing, cross-links of the immunoprecipitated chromatin were reversed.
107 DNA was recovered with the Chelex-100 protocol (16). Recovered DNA was purified with a
108 PCR purification kit (Qiagen). ChIP DNA was used to Q-PCR analysis (Supplemental Table 2)
109 and were treated at 95°C for 5 min before used for library preparation of ssDNA in accordance
110 with previously published procedures (17).

111

112 **ChIP-ssSeq sequencing and analysis**

113 The ssDNA libraries were sequenced using paired-end sequencing on Illumina Hi-Seq 2000 or
114 2500 machines. Reads were first mapped to the yeast genome (sacCer3) using Bowtie2 software
115 (18). Consistent pair-end reads were chosen for subsequent analysis. We noted that after removal
116 of duplicated reads, pair-end reads with the same ends were rarely detected in our samples even
117 for Mcm6 ChIP-seq using G1 cells. This is likely due to the fact that chromatin was sheared by
118 sonication and the ends were processed during library preparation. The genome-wide read
119 coverage of Watson and Crick strands was calculated by BEDTools (19). The reads of the
120 Watson and Crick strands were merged for peak calling by using MACS software (20).

121

122 We used our previously mapped DNA origins dataset for analysis (12). To calculate the average
123 bias pattern, the average \log_2 ratios of sequencing reads of Watson strand over Crick strand
124 surrounding 134 early replication origins (± 10 Kb and ± 30 Kb of HU-stalled and active forks,

125 respectively) were calculated using a sliding window of 200-bp. The duplicate reads were
126 excluded from calculation. These ratios were then normalized against the corresponding input to
127 obtain the average bias pattern of ChIP-ssSeq. To analyze bias at individual origins, each peak
128 region was separated into 4 quadrants: Watson strand at the left (WL) and right (WR) of an
129 origin and Crick strand at the left (CL) and right (CR) of an origin. The number of sequence
130 reads in each quadrant was counted. The binomial distribution was used to calculate the *P* value
131 to determine whether sequence reads at leading strand (WL+CR) were different from those of the
132 lagging strand (WR+CL) at each replication fork. The \log_2 ratio ($\log_2 [(WL+CR)/(WR+CL)]$) at
133 each replication origin was calculated and used to determine whether a ChIP-ssSeq peak had a
134 positive or negative strand bias.

135 Published dataset used in this study: Gene Expression Omnibus under accession number
136 GSE52614.

137

138 **Results**

139 **Rationale for analyzing replication proteins using ChIP-ssSeq**

140 ChIP-sequencing (ChIP-seq) has been widely used to study the association pattern of a protein of
141 interest with chromatin (21). Most ChIP-seq libraries are prepared using protocols that involve
142 ligation of dsDNA, which often leads to loss of ssDNA and strand-specific information (Fig. 1A).
143 During DNA replication, dsDNA is unwound to generate ssDNA, which serves as the template
144 for DNA synthesis. In the process, replication proteins, including ssDNA-binding proteins RPA
145 and DNA polymerases, may partially interact with the ssDNA or DNA-RNA hybrids. In addition,

146 DNA replication forks consist of leading and lagging strand DNA synthesis, and strand-specific
147 information helps elucidate how a protein interacts with forks.

148

149 We previously reported development of the enrichment and sequencing
150 protein-associated nascent DNA (eSPAN) method, which detects the association of a replication
151 protein with nascent leading/lagging strand DNA (Fig.1B, right panel) (12). However, this
152 method loses the information on how a protein interacts with ssDNA, which is prevalent at DNA
153 replication forks. We also generated ChIP-ssSeq datasets during the process of obtaining eSPAN
154 datasets. Briefly, protein ChIP DNA was denatured and ligated to the 3' end of an adaptor (oligo)
155 (Illumina) with an ssDNA ligase; ssDNA was then converted into dsDNA and ligated to a
156 second adaptor (17, 22). The sequence reads were mapped to the Watson and Crick strands of the
157 yeast genome (Fig. 1B). Since DNAs for a protein ChIP-ssSeq likely contain both template and
158 nascent DNA (Fig.1B, left panel), ChIP-ssSeq will allow us to deduce how a DNA replication
159 protein associates with single-stranded template DNA. As discussed and shown below, the ChIP-
160 ssSeq and the eSPAN are two complementary methods, with each revealing unique information
161 on the association of a protein at DNA replication forks.

162

163 **RPA ChIP-ssSeq shows that RPA is enriched at the lagging strand template**

164 We first analyzed Rfa1 ChIP-ssSeq datasets to gain insight into how RPA associates with DNA
165 replication forks. Briefly, yeast cells were arrested at G1 and then released into early S phase in
166 the presence of HU for 45 minutes. Rfa1 (the large subunit of the RPA complex) ChIP was

167 performed with G1 cells and early S-phase cells. Rfa1 was barely detectable at the replication
168 origin (*ARS607*) or at a distal site (*ARS607+8 kb*, unreplicated region) at G1 (Supplemental Fig.
169 1A and 1B). In contrast, Rfa1 was enriched 10-fold at *ARS607* compared with the distal site
170 (*ARS607+8 kb*) in the presence of HU (Supplemental Fig. 1B), indicating that RPA is recruited
171 to DNA replication forks during S phase. Under these conditions, replication checkpoint kinase
172 Rad53 is activated as shown by Western blot analysis of Rad53 (Supplemental Fig. 1C). In
173 addition, the fact that late origins were not fired under these conditions also reflects the
174 activation of Rad53 checkpoint kinase. Rfa1 ChIP-ssSeq peaks at *ARS510* and *ARS511* were
175 asymmetric surrounding each origin (Fig. 2A), consistent with RPA binding to ssDNA and not to
176 dsDNA. We note that a previous study shows that RPA binds asymmetrically to resected ssDNA
177 in a double-strand break site (23).

178

179 To analyze Rfa1 ChIP-ssSeq results quantitatively at a genome-wide scale, we first
180 calculated the average bias pattern, which is the average \log_2 ratio of sequencing reads of Watson
181 strand over Crick strand using 200-bp sliding window surrounding 134 early replication origins.
182 The average bias pattern of Rfa1 ChIP-ssSeq peaks indicated that on the right side of origin,
183 RPA bound more to the Watson strand, whereas on the left side of origin, it bound more to the
184 Crick strand (Fig. 2B). We categorized this finding as a positive (+) bias pattern to differentiate it
185 from the leading-strand bias pattern revealed by the eSPAN method, which detects the
186 association of a protein with newly synthesized DNA (12). As controls, the Rfa1 ChIP-ssSeq
187 using G1 cells did not show any bias (Fig. 2B), suggesting that bias seen in early S phase reflects
188 how RPA associates with DNA replication forks in the presence of HU. We also analyzed the

189 bias pattern of Rfa1 ChIP-ssSeq peaks at each of the 134 individual replication origins (Fig. 2C).
190 Rfa1 ChIP-ssSeq peaks showed (+) bias for most origins (n=89 [66%]), whereas the Rfa1 ChIP-
191 ssSeq using G1 phase cells showed no bias for the majority of origins (n=119 [89%]). These
192 results support the idea that RPA binds ssDNA of DNA replication forks stalled by HU.

193

194 While RPA is known to bind single-stranded template DNA, it may also contact nascent
195 DNA at replication forks indirectly through protein-protein interactions. Indeed, RPA eSPAN
196 reveals that RPA bind more to nascent lagging strands (12). Rfa1 ChIP-DNA contains the
197 template strand and the nascent strand DNA. Two potential mechanisms account for the (+) bias
198 pattern of Rfa1 ChIP-ssSeq peaks. First, (+) bias may indicate that more RPA binds to the
199 lagging strand template than to the corresponding leading strand template (Fig. 2D). Second,
200 RPA may bind more nascent leading strands than the corresponding nascent lagging strands.
201 However, the later explanation contradicts the RPA eSPAN results outlined above (12) (Fig 2D).
202 Based on our Rfa1 ChIP-ssSeq and Rfa1 eSPAN results, we suggest that more RPA binds
203 lagging strand template than leading strand template of HU-stalled forks (Fig. 2D). The above
204 RPA ChIP experiment is under HU condition. We also performed the RPA ChIP-ssSeq under
205 normal condition. The results showed the same (+) bias pattern (Supplemental Fig. 1D-E),
206 suggesting that more RPA are enriched at lagging strand template compared to leading strand
207 template at both active and HU stalled forks. This explanation is consistent with the proposed
208 model of RPA preferentially binding the lagging template strand to protect gaps between
209 Okazaki fragments (24). To our knowledge, the result is the first experimental demonstration that
210 more RPA binds lagging strand template than leading strand template. In addition to DNA

211 replication, RPA is also involved in DNA repair process and activation of DNA replication
212 checkpoint (25, 26).

213

214 **PCNA ChIP-ssSeq shows no strand bias at replication forks**

215 We analyzed PCNA ChIP-ssSeq datasets obtained from cells cultured with or without HU. No
216 obvious strand bias was observed from the analysis of the average bias pattern of all early
217 replication origins or the analysis of the bias pattern of individual origins (Supplemental Fig. 2A-
218 C). These results indicate that PCNA, a processivity factor of DNA polymerases that is loaded
219 onto primer-template junctions, contacts dsDNA including both template and nascent DNA at
220 active and HU-stalled replication forks.

221

222 **MCM ChIP-ssSeq shows no strand bias at stalled replication forks**

223 Analysis of Mcm6 ChIP-ssSeq showed no significant strand bias of HU-stalled forks
224 (Supplemental Fig. 2D-F), suggesting that the MCM helicase binds to dsDNA. Similar results
225 were obtained for Mcm4 ChIP-ssSeq (Supplemental Fig. 2D-F). This observation seems to
226 contradict the idea that the MCM complex travels along the leading strand (12, 27), and our
227 eSPAN results showing that MCM associates preferentially with nascent leading strand DNA
228 compared with nascent lagging strand DNA. One likely explanation for our ChIP-ssSeq results is
229 that the MCM helicase, while encircling one leading template DNA strand, still makes indirect
230 contact with another lagging template strand of HU-stalled forks. Indeed, it has been shown that
231 MCM protein complex interacts with both Pol ϵ , which is enriched at leading strand, and Pol α ,

232 which is enriched at lagging strands (28-31). For the rest of our studies, we focused on analysis
233 on how three DNA polymerases associate with active and HU-stalled replication forks.

234

235 **Pol α ChIP-ssSeq indicates that Pol α preferentially binds to DNA-RNA hybrids at lagging**
236 **strands of active and HU-stalled replication forks**

237 Pol α was enriched at the early replication origin (*ARS607*) compared with the distal site
238 (*ARS607*+8 kb) when cells were released from G1 to early S phase in the presence of HU
239 (Supplemental Fig. 3A-B), consistent with the results that Pol α associates with replicating DNA
240 even in the presence of HU (32, 33). Inspection of Pol α ChIP-ssSeq at replication origins
241 *ARS510* and *ARS511* showed that Pol α ChIP-ssSeq peaks showed a strong (+) bias pattern
242 (Fig.3A and 3B). Analysis of the average bias pattern of 134 early replication origins confirmed
243 that the (+) bias pattern of Pol α ChIP-ssSeq peaks at individual forks on a genome-wide scale
244 (Fig. 3B), with 126 of 134 peaks (94%) showing (+) bias (Fig. 3C). We also determined how Pol
245 α bound to active replication forks by performing ChIP-ssSeq using cells released into S phase
246 without HU at a lower temperature (Fig. 3C and Supplemental Fig. 3C). The Pol α ChIP-ssSeq
247 peaks also showed (+) bias based on the analysis of average bias pattern of early replication
248 origins, as well as in the analysis of individual origins (Fig. 3C-E).

249 Pol α ChIPed DNA consists of the lagging strand template and newly synthesized RNA-
250 DNA primer. The (+) bias pattern of Pol α ChIP-ssSeq peaks indicates that more Pol α binds to
251 lagging strand template than to leading strand template of active and HU-stalled replication forks.
252 Supporting this idea, the published Pol α eSPAN peaks indicate that Pol α physically binds more

253 to nascent lagging strands than to leading strands at active and HU-stalled replication forks (12)
254 (Fig. 3B, E-F). Because Pol α is involved in the synthesis of RNA and DNA primers, the (+) bias
255 indicates that Pol α binds to initiating Okazaki fragments at HU-stalled forks and to the
256 elongating Okazaki fragments of active replication forks (Fig. 3F).

257

258 **Pol δ is enriched at elongating Okazaki fragments of lagging strand template only of active**
259 **replication forks**

260 We next analyzed Pol δ (catalytic subunit) ChIP-ssSeq obtained using cells released from G1
261 arrest into early S phase in the presence of HU. Pol δ ChIP-PCR analysis showed that Pol δ was
262 enriched at replication forks originating from *ARS607* comparing to the unreplicated distal site
263 (*ARS607*+8 kb) (Supplemental Fig. 4A-B). Pol δ ChIP-ssSeq peaks at HU-stalled forks did not
264 reveal any bias pattern based on analysis of the average bias of 134 forks from early replication
265 origins or with analysis of individual origins (Fig. 4A-C), suggesting that Pol δ binds equally to
266 Watson and Crick strands of HU-stalled replication forks. In contrast, Pol δ ChIP-ssSeq peaks at
267 active replication forks without HU showed small, but consistent (+) bias in the analysis of the
268 average bias pattern and of individual forks at all three-time points (Fig. 4B-C). The difference in
269 Pol δ ChIP-ssSeq peak bias between active and HU-stalled forks was unlikely due to difference
270 in input samples (Supplemental Fig. 4C). Thus, Pol δ differentially associates with DNA at
271 active and HU-stalled replication forks

272

273 In principle, Pol δ binds both template and nascent DNA. Therefore, Pol δ ChIP-ssSeq peaks
274 should show no bias at both active and HU-stalled forks. The eSPAN analysis of Pol δ indicates
275 that Pol δ binds preferentially nascent lagging strand of HU-stalled and active replication forks
276 (12) (Fig. 4B). We therefore deduce from the (+) bias pattern of Pol δ ChIP-ssSeq peaks that Pol
277 δ associates with more lagging strand template of active forks, which most likely reflect that Pol
278 δ can associates with newly initiated Okazaki fragments with only very short nascent RNA of
279 active forks (Fig. 4D-E). In contrast, this mode of association of Pol δ is lost at HU-stalled forks,
280 which provides an explanation for a lacking of bias of Pol δ ChIP-ssSeq peaks of HU-stalled
281 forks. We have shown recently that the DNA polymerase clamp, PCNA, is unloaded from
282 lagging strands of HU-stalled forks (12). PCNA is important for the activity of Pol δ , likely
283 important for tethering Pol δ at DNA replication forks. Therefore, the unloading of PCNA from
284 lagging strand of HU stalled forks may contribute to the loss of the association of Pol δ newly
285 initiated Okazaki fragment at HU-stalled forks, whereas Pol α still binds.

286

287 **Pol ϵ -DNA interaction is different for active and the HU-stalled replication forks.**

288 After determining the association of Pol α and Pol δ with DNA, we next used ChIP-ssSeq to
289 examine how Pol ϵ interacts with DNA. Pol ϵ ChIP-PCR analysis indicated that Pol ϵ bind to
290 replicating DNA at HU-stalled replication forks (Supplemental Fig. 4A-B). Like Pol δ , Pol ϵ
291 ChIP-ssSeq showed that Pol ϵ did not show significant bias at HU-stalled replication forks from
292 early replication origins (Fig. 5A-C), indicating that Pol ϵ is cross-linked to dsDNA, including
293 the template strand and nascent leading strand of HU-stalled replication forks. Remarkably, Pol ϵ

294 ChIP-ssSeq showed (+) bias at actively replicating forks at all time points considered (72, 84,
295 and 96 minutes after release from G1) (Fig. 5B). The bias pattern was detected at the majority of
296 individual origins (Fig. 5C), suggesting that the Pol ϵ -DNA interaction at active forks differs
297 from that at stalled forks.

298

299 The above Pol ϵ ChIP-ssSeq analysis of HU-stalled and active replication forks were
300 obtained from independent experiments and Pol ϵ ChIP-ssSeq bias is small. Therefore, we
301 performed additional experiments to confirm that different association patterns of Pol ϵ with
302 DNA changes in stalled vs active forks. Briefly, yeast cells were arrested in G1 with α factor and
303 then released into HU-containing medium for 45 minutes. A fraction of cells were collected for
304 Pol ϵ ChIP-ssSeq, and the remaining cells were released into fresh medium without HU. Samples
305 were used to perform Pol ϵ ChIP-ssSeq at 3 time points after release from HU (20, 30, and 40
306 minutes) (Fig. 5D). Analysis of Pol ϵ ChIP-ssSeq datasets showed no bias pattern for peaks
307 obtained using cells treated with HU, whereas peaks from cells after HU removal showed (+)
308 bias (Fig 5E-G). We noticed that Pol ϵ ChIP-ssSeq at HU conditions shown in Fig. 5B and 5F
309 appears to show opposite trend. This is likely due to the fact that Pol ϵ ChIP-ssSeq peaks at most
310 origins showed indeterminable bias (no bias) and variations at a small number of origins
311 contributes to the apparent changes in the insignificant bias pattern (compare Fig. 5C and Fig.
312 5G). Nonetheless, we observed very consistent results of Pol ϵ ChIP-ssSeq at active forks from
313 each of the 3 time points of two independent experiments, supporting the idea that the
314 association of Pol ϵ with DNA is altered when active forks become stall by HU treatment.

315 Once again, two potential models explain the (+) bias pattern of Pol ϵ CHIP-ssSeq peaks
316 (Fig. 5H). Based on Pol ϵ eSPAN results (12) (Fig. 5B), Pol ϵ binds preferentially to leading
317 strand. Therefore, it is possible that in addition to contact with leading strand DNA, Pol ϵ may
318 also directly contact the lagging strand template during normal replication. This mechanism is
319 unlikely because it is hard to put the Cdc45-MCM-GINS complex, which is known to associate
320 with Pol ϵ on the leading strand (29), in front of Pol ϵ . Second, Pol ϵ may not contact the leading
321 strand template tightly, binding only to nascent DNA on the leading strand of active forks (Fig.
322 5H). We suggest that this mode of interaction with leading nascent DNA facilitates its ability to
323 proofread or repair mis-incorporated nucleotides by using its 3'-to-5' exonuclease activity (34).
324 At stalled fork, Pol ϵ may backtrack and associate with dsDNA including both template and
325 nascent strands (Fig. 5I).

326

327 **Discussion**

328 Our present study reveals several novel insights into the contacts of proteins with active and HU-
329 stalled forks. First, we provide the experimental evidence that RPA are enriched at lagging strand
330 template compared to the corresponding leading strand template, consistent with the replication
331 model on the role of RPA in DNA replication. Second, we show that Pol α associates with
332 lagging strand template of both active and HU-stalled forks. Third, we show that both Pol δ and
333 Pol ϵ bind to HU-stalled forks differently from active forks. Specifically, Pol δ binds to both
334 initiating and elongating Okazaki fragments at active forks, and is likely lost/removed from
335 initiating Okazaki fragments at HU-stalled forks where Pol α remains to be present. Pol ϵ likely

336 backtracks and associates dsDNA at HU-stalled forks. These results provide insight into how
337 DNA synthesis can resume soon after removal of HU-induced replication stress.

338

339 **Advantages and limitations of ChIP-ssSeq method and its comparison with the eSPAN**
340 **method**

341 The library preparation of traditional ChIP-Seq includes steps for dsDNA repair and dsDNA
342 ligation. During the sample preparation process, protein-bound ssDNA and strand-specific
343 information is lost (Fig. 1A). Generally, this loss is not an issue because most proteins bind
344 dsDNA. However, during DNA replication, dsDNA is transiently unwound into ssDNA.
345 Therefore, determining whether a DNA replication protein binds to ssDNA will help elucidate its
346 mode of action in DNA synthesis. Here, we used ChIP-ssSeq to gain insights into how DNA
347 replication proteins bind to active and stalled replication forks. Since DNAs for protein ChIP-
348 ssSeq potentially contain both template and nascent DNA, we analyzed ChIP-ssSeq peaks by
349 calculating the average log₂ ratio of sequence reads of Watson over Crick strand. If a protein
350 contacts dsDNA including both template and nascent DNA, the ratio should be zero without any
351 bias towards Watson or Crick strand. If a protein binds ssDNA, the average log₂ ratio of
352 sequence reads of Watson over Crick strand of ChIP-ssSeq peaks is not zero. Indeed, ssDNA-
353 binding protein RPA ChIP-ssSeq peaks at both active and HU-stalled forks exhibit (+) bias,
354 indicating that more RPA are present at lagging strand template than leading strand template,
355 consistent with replication models of RPA in DNA replication.

356

357 We reported previously the development of the eSPAN method, which detects how a
358 protein associates with newly synthesized DNA at DNA replication forks (12). Using this
359 method, we detect the association of different DNA replication proteins with nascent leading or
360 lagging strand of DNA replication forks. For instance, we observed that Pole and Pol δ are
361 enriched at nascent leading and lagging strand, respectively, consistent with their division of
362 labor during DNA synthesis. Interestingly, we also show that RPA is enriched at lagging strand
363 using eSPAN. This result appears to contradict the idea that RPA binds and stabilizes single-
364 stranded template DNA. Using ChIP-ssSeq, we show that RPA is enriched at lagging strand
365 template compared to leading strand template. One explanation for the apparent discrepancy for
366 RPA ChIP-ssSeq and eSPAN results is that RPA binds preferentially to lagging strand template,
367 but contacts nascent DNA indirectly, likely through other proteins. In this way, RPA eSPAN
368 peaks show lagging strand bias. Similarly, the PCNA eSPAN results show that PCNA is
369 enriched at nascent lagging strand DNA of active forks and nascent leading strand DNA at HU-
370 stalled forks, suggesting that PCNA is unloaded from lagging strand of HU-stalled forks (12).
371 However, PCNA ChIP-ssSeq peaks at HU-stalled forks show no bias pattern. One explanation is
372 that PCNA contacts both template DNA and nascent DNA at HU-stalled forks, which gives rise
373 to the no bias pattern of PCNA ChIP-ssSeq peaks. Therefore, the bias pattern of eSPAN peaks
374 and ChIP-ssSeq peaks has a different meaning: the eSPAN peak bias indicates how a protein
375 associates with nascent leading and lagging strands of DNA replication forks, whereas ChIP-
376 ssSeq peak bias reflects how a protein binds to ssDNA and dsDNA. Together, these results
377 comparing ChIP-ssSeq and eSPAN of different proteins including RPA, PCNA and MCM,
378 PCNA and DNA polymerases (below) indicate that the ChIP-ssSeq and eSPAN methods provide
379 complementary information on how a protein associates with DNA replication forks.

380

381 In theory, ChIP-ssSeq is suitable for studying DNA repair and RNA transcription when
382 strand-specific information is needed. In fact, some reports indicated that similar ChIP-ssSeq
383 approaches can study the DNA repair process (23, 35). These studies used either sticky end
384 dsDNA adaptor ligation or intramolecular microhomology to generate libraries. We adopted the
385 ssDNA library preparation method developed by Meyer et al (17), which was used to analyze
386 highly damaged DNA from ancient human samples. The advantage of this method, compared
387 with the 2 published ssDNA library preparation methods, is high efficiency. Meyer's method can
388 generate libraries from very low quantities of DNA (22), and it therefore is very suitable for
389 constructing libraries from the low amount of DNA isolated by ChIP experiments. We expect
390 that ChIP-ssSeq may also yield useful information for other processes besides DNA replication
391 and repair. For instance, allele-specific DNA methylation is known to occur frequently in
392 mammalian cells (36). A combination of library preparation methods with immunoprecipitation
393 of methylated DNA will, in principle, be able to differentiate between methylated and
394 unmethylated alleles. Future studies are needed to test this idea.

395

396 **Association of Pol α with active and HU-stalled replication forks**

397 The Pol α -primase complex synthesizes primers for subsequent DNA synthesis by Pol δ and Pol
398 ϵ , likely at the lagging and leading strands of DNA replication forks, respectively. Previously, we
399 used the eSPAN method to show that Pol α -primase is enriched at the nascent lagging strand of
400 the DNA replication fork; this finding was consistent with the classical replication models that

401 require a Pol α -primase complex for each Okazaki fragment (37, 38). In this study, we used the
402 Pol α ChIP-ssSeq method to show that Pol α also binds more to the lagging strand template
403 during DNA replication, further supporting the role of Pol α -primase in the classical DNA
404 replication model. Interestingly, we observed that Pol α -primase remains binding to the lagging
405 strand template at HU-stalled replication forks. We suggest that the association of Pol α -primase
406 complex with the template strand under this condition may facilitate resumption of the
407 replication process soon after amelioration of DNA replication stress. In addition, this
408 association may serve as the target of cell cycle checkpoint kinases that regulate arrest of DNA
409 replication forks during stress. Consistent with this idea, previous work has shown that primase
410 connects DNA replication to the DNA damage response (39).

411

412 **Altered association of Pol δ and Pol ϵ with DNA replication forks stalled by replication**
413 **stress**

414 Pol δ replicates both leading and lagging strands in the SV40 *in vitro* replication system (40-42).
415 However, genetic evidence from the past decade supports Pol ϵ as the leading-strand replication
416 enzyme in yeast (5, 6, 8, 43, 44) and Pol δ is responsible for replicating lagging-strand DNA.
417 Recently, the division of labor between Pol ϵ and Pol δ in DNA synthesis has come into question
418 with genetic analyses of mismatch repair-deficient DNA polymerase mutants (13). Therefore,
419 several eukaryotic DNA replication models have been proposed (45). In every model, Pol ϵ is
420 always physically linked with MCM helicase on the leading strand, regardless of whether it is the
421 major active leading-strand DNA polymerase or just a repair enzyme. The result is fully
422 compatible with our eSPAN data, indicating that Pol ϵ is enriched at the replicating leading

423 strands. In contrast, Pol δ is enriched at the nascent lagging strands of DNA replication forks.

424 We show here that Pol δ and Pol ϵ asymmetrically bind to DNA of active replication forks,

425 suggesting that these two polymerases also bind to ssDNA, but not solely to dsDNA at active

426 replication forks. In contrast, at HU-stalled forks, Pol δ and Pol ϵ predominantly were bound to

427 dsDNA. We suggest that at HU-stalled forks, Pol ϵ may backtrack to contact dsDNA.

428

429 Previous study has shown that MCM localization can be displaced several hundred base pairs

430 from the origin by transcription regulation (46). While it is possible that transcriptional alteration

431 during HU block contributes to the lack of bias Pol ϵ and Pol δ ChIP-ssSeq peaks at HU-stalled

432 forks, it is unlikely for the following reasons, First, we show that the Pol ϵ ChIP-ssSeq peak bias

433 pattern reappears after we release cells from HU block to fresh media, suggesting that Pol ϵ bias

434 is associated with active replication forks. Second, it is known that HU has no apparent effect on

435 initiation of DNA replication at early replication origins based on studies from many laboratories.

436 Moreover, the observation that transcription can shift MCM localization was made in *rat1*

437 mutant cells in which transcription termination was reduced, whereas at HU-stalled forks, we did

438 not observe such dramatic alterations in MCM distribution (12).

439

440 We noticed the bias of Pol δ and Pol ϵ ChIP-ssSeq peaks at active forks is small compared to that

441 of Rfa1 or Pol α . The small bias is not likely an artifact of calculation because we analyzed ChIP-

442 ssSeq data sets using two different methods. First, we used 200 bp sliding window to calculate

443 bias from either 10 or 30Kb surrounding each replication origins of HU-stalled and active forks,

444 respectively. The trend of each data point of Pol δ and Pol ϵ CHIP-ssSeq show that the bias, while
445 small, is not random. Second, we also analyzed whether there exist bias of Pol δ and Pol ϵ CHIP-
446 ssSeq peaks at individual replication forks (Fig. 4C, Fig. 5C and Fig. 5G) and found that the bias,
447 while small, is statistically significant. Unlike RPA that binds ssDNA and Pol α that synthesizes
448 primers for Pol δ and Pol ϵ , most Pol δ and Pol ϵ likely still contact dsDNA including template
449 DNA and newly synthesized DNA. Therefore, it is not surprising that Pol δ and Pol ϵ CHIP-ssSeq
450 bias is smaller than RPA or Pol α . In conclusion, the bias changes at HU-stalled forks and active
451 forks of Pol δ and Pol ϵ CHIP-ssSeq, while small, reflect the polymerase-DNA spatial contacts at
452 active forks.

453

454 **Acknowledgement**

455 We thank Dr. Oscar Aparicio for yeast strains and plasmids. We thank Dr. Steven Brill for the
456 anti-RPA antibody and Dr. June Oshiro for editing our manuscript.

457

458

459 **Competing financial interests**

460 The authors declare no competing financial interests

461

462

463 References

- 464 1. **Zeman MK, Cimprich KA.** 2014. Causes and consequences of replication stress. *Nat Cell Biol*
465 **16**:2-9.
- 466 2. **Mazouzi A, Velimezi G, Loizou JI.** 2014. DNA replication stress: causes, resolution and disease.
467 *Exp Cell Res* **329**:85-93.
- 468 3. **Aparicio OM.** 2013. Location, location, location: it's all in the timing for replication origins.
469 *Genes Dev* **27**:117-128.
- 470 4. **Bell SD, Botchan MR.** 2013. The minichromosome maintenance replicative helicase. *Cold Spring*
471 *Harb Perspect Biol* **5**:a012807.
- 472 5. **Pursell ZF, Isov I, Lundstrom EB, Johansson E, Kunkel TA.** 2007. Yeast DNA polymerase epsilon
473 participates in leading-strand DNA replication. *Science* **317**:127-130.
- 474 6. **Nick McElhinny SA, Gordenin DA, Stith CM, Burgers PM, Kunkel TA.** 2008. Division of labor at
475 the eukaryotic replication fork. *Mol Cell* **30**:137-144.
- 476 7. **Stillman B.** 2008. DNA polymerases at the replication fork in eukaryotes. *Mol Cell* **30**:259-260.
- 477 8. **Clausen AR, Lujan SA, Burkholder AB, Orebaugh CD, Williams JS, Clausen MF, Malc EP,**
478 **Mieczkowski PA, Fargo DC, Smith DJ, Kunkel TA.** 2015. Tracking replication enzymology in vivo
479 by genome-wide mapping of ribonucleotide incorporation. *Nat Struct Mol Biol* **22**:185-191.
- 480 9. **Daigaku Y, Keszthelyi A, Muller CA, Miyabe I, Brooks T, Retkute R, Hubank M, Nieduszynski CA,**
481 **Carr AM.** 2015. A global profile of replicative polymerase usage. *Nat Struct Mol Biol* **22**:192-198.
- 482 10. **Reijns MA, Kemp H, Ding J, de Proce SM, Jackson AP, Taylor MS.** 2015. Lagging-strand
483 replication shapes the mutational landscape of the genome. *Nature* **518**:502-506.
- 484 11. **Koh KD, Balachander S, Hesselberth JR, Storic F.** 2015. Ribose-seq: global mapping of
485 ribonucleotides embedded in genomic DNA. *Nat Methods* **12**:251-257, 253 p following 257.
- 486 12. **Yu C, Gan H, Han J, Zhou ZX, Jia S, Chabes A, Farrugia G, Ordog T, Zhang Z.** 2014. Strand-Specific
487 Analysis Shows Protein Binding at Replication Forks and PCNA Unloading from Lagging Strands
488 when Forks Stall. *Mol Cell* **56**:551-563.
- 489 13. **Johnson RE, Klassen R, Prakash L, Prakash S.** 2015. A Major Role of DNA Polymerase delta in
490 Replication of Both the Leading and Lagging DNA Strands. *Mol Cell* **59**:163-175.
- 491 14. **Santocanale C, Diffley JF.** 1998. A Mec1- and Rad53-dependent checkpoint controls late-firing
492 origins of DNA replication. *Nature* **395**:615-618.
- 493 15. **Shirahige K, Hori Y, Shiraishi K, Yamashita M, Takahashi K, Obuse C, Tsurimoto T, Yoshikawa H.**
494 1998. Regulation of DNA-replication origins during cell-cycle progression. *Nature* **395**:618-621.
- 495 16. **Nelson JD, Denisenko O, Bomsztyk K.** 2006. Protocol for the fast chromatin
496 immunoprecipitation (ChIP) method. *Nat Protoc* **1**:179-185.
- 497 17. **Meyer M, Kircher M, Gansauge MT, Li H, Racimo F, Mallick S, Schraiber JG, Jay F, Prufer K, de**
498 **Filippo C, Sudmant PH, Alkan C, Fu Q, Do R, Rohland N, Tandon A, Siebauer M, Green RE, Bryc**
499 **K, Briggs AW, Stenzel U, Dabney J, Shendure J, Kitzman J, Hammer MF, Shunkov MV,**
500 **Derevianko AP, Patterson N, Andres AM, Eichler EE, Slatkin M, Reich D, Kelso J, Paabo S.** 2012.
501 A high-coverage genome sequence from an archaic Denisovan individual. *Science* **338**:222-226.
- 502 18. **Langmead B, Salzberg SL.** 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**:357-
503 359.
- 504 19. **Quinlan AR, Hall IM.** 2010. BEDTools: a flexible suite of utilities for comparing genomic features.
505 *Bioinformatics* **26**:841-842.
- 506 20. **Zhang Y, Liu T, Meyer CA, Eeckhoutte J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM,**
507 **Brown M, Li W, Liu XS.** 2008. Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**:R137.

- 508 21. **Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K.** 2007.
509 High-resolution profiling of histone methylations in the human genome. *Cell* **129**:823-837.
- 510 22. **Gansauge MT, Meyer M.** 2013. Single-stranded DNA library preparation for the sequencing of
511 ancient or damaged DNA. *Nat Protoc* **8**:737-748.
- 512 23. **Yamane A, Robbiani DF, Resch W, Bothmer A, Nakahashi H, Oliveira T, Rommel PC, Brown EJ,**
513 **Nussenzweig A, Nussenzweig MC, Casellas R.** 2013. RPA accumulation during class switch
514 recombination represents 5'-3' DNA-end resection during the S-G2/M phase of the cell cycle.
515 *Cell Rep* **3**:138-147.
- 516 24. **O'Donnell M, Langston L, Stillman B.** 2013. Principles and concepts of DNA replication in
517 bacteria, archaea, and eukarya. *Cold Spring Harb Perspect Biol* **5**.
- 518 25. **Marechal A, Zou L.** 2013. DNA damage sensing by the ATM and ATR kinases. *Cold Spring Harb*
519 *Perspect Biol* **5**.
- 520 26. **Zou L, Elledge SJ.** 2003. Sensing DNA damage through ATRIP recognition of RPA-ssDNA
521 complexes. *Science* **300**:1542-1548.
- 522 27. **Fu YV, Yardimci H, Long DT, Ho TV, Guainazzi A, Bermudez VP, Hurwitz J, van Oijen A, Scharer**
523 **OD, Walter JC.** 2011. Selective bypass of a lagging strand roadblock by the eukaryotic replicative
524 DNA helicase. *Cell* **146**:931-941.
- 525 28. **Langston LD, Zhang D, Yurieva O, Georgescu RE, Finkelstein J, Yao NY, Indiani C, O'Donnell ME.**
526 2014. CMG helicase and DNA polymerase epsilon form a functional 15-subunit holoenzyme for
527 eukaryotic leading-strand DNA replication. *Proc Natl Acad Sci U S A* **111**:15390-15395.
- 528 29. **Sengupta S, van Deursen F, de Piccoli G, Labib K.** 2013. Dpb2 integrates the leading-strand DNA
529 polymerase into the eukaryotic replisome. *Curr Biol* **23**:543-552.
- 530 30. **Gambus A, Jones RC, Sanchez-Diaz A, Kanemaki M, van Deursen F, Edmondson RD, Labib K.**
531 2006. GINS maintains association of Cdc45 with MCM in replisome progression complexes at
532 eukaryotic DNA replication forks. *Nat Cell Biol* **8**:358-366.
- 533 31. **Gambus A, van Deursen F, Polychronopoulos D, Foltman M, Jones RC, Edmondson RD, Calzada**
534 **A, Labib K.** 2009. A key role for Ctf4 in coupling the MCM2-7 helicase to DNA polymerase alpha
535 within the eukaryotic replisome. *EMBO J* **28**:2992-3004.
- 536 32. **Cobb JA, Bjergbaek L, Shimada K, Frei C, Gasser SM.** 2003. DNA polymerase stabilization at
537 stalled replication forks requires Mec1 and the RecQ helicase Sgs1. *Embo Journal* **22**:4325-4336.
- 538 33. **De Piccoli G, Katou Y, Itoh T, Nakato R, Shirahige K, Labib K.** 2012. Replisome stability at
539 defective DNA replication forks is independent of S phase checkpoint kinases. *Mol Cell* **45**:696-
540 704.
- 541 34. **Tran HT, Gordenin DA, Resnick MA.** 1999. The 3'-->5' exonucleases of DNA polymerases delta
542 and epsilon and the 5'-->3' exonuclease Exo1 have major roles in postreplication mutation
543 avoidance in *Saccharomyces cerevisiae*. *Mol Cell Biol* **19**:2000-2007.
- 544 35. **Zhou ZX, Zhang MJ, Peng X, Takayama Y, Xu XY, Huang LZ, Du LL.** 2013. Mapping genomic
545 hotspots of DNA damage by a single-strand-DNA-compatible and strand-specific ChIP-seq
546 method. *Genome Res* **23**:705-715.
- 547 36. **Kerkel K, Spadola A, Yuan E, Kosek J, Jiang L, Hod E, Li K, Murty VV, Schupf N, Vilain E, Morris**
548 **M, Haghghi F, Tycko B.** 2008. Genomic surveys by methylation-sensitive SNP analysis identify
549 sequence-dependent allele-specific DNA methylation. *Nat Genet* **40**:904-908.
- 550 37. **Waga S, Stillman B.** 1998. The DNA replication fork in eukaryotic cells. *Annu Rev Biochem*
551 **67**:721-751.
- 552 38. **Bell SP, Dutta A.** 2002. DNA replication in eukaryotic cells. *Annu Rev Biochem* **71**:333-374.
- 553 39. **Marini F, Pellicoli A, Paciotti V, Lucchini G, Plevani P, Stern DF, Foiani M.** 1997. A role for DNA
554 primase in coupling DNA replication to DNA damage response. *EMBO J* **16**:639-650.

- 555 40. **Tsurimoto T, Stillman B.** 1991. Replication factors required for SV40 DNA replication in vitro. II.
556 Switching of DNA polymerase alpha and delta during initiation of leading and lagging strand
557 synthesis. *J Biol Chem* **266**:1961-1968.
- 558 41. **Tsurimoto T, Stillman B.** 1991. Replication factors required for SV40 DNA replication in vitro. I.
559 DNA structure-specific recognition of a primer-template junction by eukaryotic DNA
560 polymerases and their accessory proteins. *J Biol Chem* **266**:1950-1960.
- 561 42. **Waga S, Stillman B.** 1994. Anatomy of a DNA replication fork revealed by reconstitution of SV40
562 DNA replication in vitro. *Nature* **369**:207-212.
- 563 43. **Georgescu RE, Langston L, Yao NY, Yurieva O, Zhang D, Finkelstein J, Agarwal T, O'Donnell ME.**
564 2014. Mechanism of asymmetric polymerase assembly at the eukaryotic replication fork. *Nat*
565 *Struct Mol Biol* **21**:664-670.
- 566 44. **Williams JS, Clausen AR, Lujan SA, Marjavaara L, Clark AB, Burgers PM, Chabes A, Kunkel TA.**
567 2015. Evidence that processing of ribonucleotides in DNA by topoisomerase 1 is leading-strand
568 specific. *Nat Struct Mol Biol* **22**:291-297.
- 569 45. **Stillman B.** 2015. Reconsidering DNA Polymerases at the Replication Fork in Eukaryotes. *Mol Cell*
570 **59**:139-141.
- 571 46. **Gros J, Kumar C, Lynch G, Yadav T, Whitehouse I, Remus D.** 2015. Post-licensing Specification of
572 Eukaryotic Replication Origins by Facilitated Mcm2-7 Sliding along DNA. *Mol Cell* **60**:797-807.

573

574

575 **Legends**

576 **Figure 1. Schematic of the ChIP-ssSeq method used to detect the association of a protein**

577 **with ssDNA. (A)** The standard ChIP-seq process does not detect protein-ssDNA interactions.

578 Here we use a protein that binds to both dsDNA and ssDNA as an example to elucidate the

579 process. The upper shows a replication bubble. The first step is the protein ChIP. To prepare a

580 library for sequencing, ChIPed DNA is extracted. After DNA end repair, only dsDNA is ligated

581 with adaptors. Therefore, sequence reads contain location information for protein-dsDNA

582 interactions. The target protein is shown as a gray ball. Black peak represent the DNA location.

583 **(B)** The ChIP-ssSeq procedure preserves strand-specific information. The replication protein

584 ChIP process is the same as the standard ChIP-seq. ChIP-ssSeq library preparation utilized a

585 ssDNA ligase to ligate denatured ChIP ssDNA to a 3' adaptor, which marks the same end of each

586 ssDNA molecule. The second DNA strand is synthesized and extended with a 3' complementary

587 oligo. After end repair, the 5' end is ligated to a 5' end adaptor with T4 DNA ligase. After

588 sequencing, the reads are mapped to the Watson and Crick strands of the yeast genome to

589 determine the location of the target protein and strand-specific information. The red and green

590 lines represent the Watson and Crick strands, respectively. ChIP-seq indicates chromatin

591 immunoprecipitation and sequencing; ChIP-ssSeq, chromatin immunoprecipitation and strand-

592 specific sequencing, dsDNA, double-stranded DNA; ssDNA, single-stranded DNA; * represents

593 nucleotide analog BrdU, which is incorporated into nascent DNA during DNA replication. As a

594 comparison, the outline for eSPAN procedures is shown in B. The eSPAN procedure involves

595 immunoprecipitation of protein-associated newly synthesized DNA marked with BrdU using

596 antibodies against BrdU. Therefore, eSPAN detects the association of a protein with nascent

597 DNA at DNA replication forks.

598

599 **Figure 2. RPA preferentially binds to ssDNA on the lagging strand template. (A)** A snapshot
600 of RPA ChIP-ssSeq peaks at 2 early replication origins, *ARS510* and *ARS511*. Log-phase cells
601 were synchronized to G1 with α factor and then released into 0.2M HU for 45 minutes. G1 and
602 early S-phase cells were collected for RPA ChIP-ssSeq. The RPA ChIP was performed using
603 antibodies against the FLAG epitope, which was fused to the C-terminus of Rfa1 (the large
604 subunit of the RPA complex). Similar results were obtained by using an Rfa1-specific antibody
605 (data not shown). The red and green regions represent normalized sequence read density of the
606 Watson and Crick strands, respectively. Rfa1 eSPAN peaks are included for comparison (12). **(B)**
607 The average pattern of Rfa1 ChIP-ssSeq peaks shows positive bias. The average \log_2 ratios of
608 sequence reads of Watson and Crick strands at 134 early replication origins were calculated
609 using a 200-bp sliding window and then normalized against input values to obtain the average
610 bias pattern (blue). The lagging strand bias pattern of Rfa1 eSPAN peaks (12) (red) was used for
611 comparison. **(C)** The dot-and-box plot shows the bias pattern of Rfa1 ChIP-ssSeq peaks at
612 individual origins. The ratio of sequence reads of lagging to leading strands was calculated at
613 each of 134 early replication origins; each dot represents one origin. The colors represent 3 bias
614 patterns: red for positive bias (+); green for negative bias (-); and blue for indeterminable. **(D)**
615 Schematic of a mechanism that accounts for the positive bias pattern of RPA based on ChIP-
616 ssSeq analysis and the lagging strand bias pattern based on eSPAN analysis. The yellow ball
617 represents the RPA complex. * represents nucleotide analog BrdU, which is incorporated into
618 nascent DNA during DNA replication. Please note that the eSPAN peak bias reflects whether a
619 protein binds to nascent leading and lagging strand, whereas ChIP-ssSeq peak bias indicates that
620 whether a protein binds to ssDNA or dsDNA.

621

622 **Figure 3. Pol α preferentially binds the single-stranded lagging template of both active and**
623 **HU-stalled replication forks. (A-C)** Pol α ChIP-ssSeq peaks at HU-stalled forks exhibit a
624 positive (+) bias pattern. **(A)** Snapshot of Pol α ChIP-ssSeq and eSPAN peaks at *ARS510* and
625 *ARS511*. The signals represent normalized sequence read density. The red and green represents
626 the Watson and Crick strands, respectively. **(B)** Analysis of the average bias pattern of Pol α
627 ChIP-ssSeq. Pol α eSPAN was used for comparison. **(C)** Analysis of bias pattern of Pol α ChIP-
628 ssSeq of HU-stalled and active replication forks at individual origins (early replication origins
629 only). **(D-E)** Pol α ChIP-ssSeq peaks at active forks show positive strand bias. G1-synchronized
630 yeast cells were released into fresh medium at 16°C in the presence of BrdU for 72 minutes. **(D)**
631 Snapshot of Pol α ChIP-ssSeq and eSPAN peaks at *ARS510* and *ARS511*. **(E)** Analysis of the
632 average bias of Pol α ChIP-ssSeq peaks at active forks. f, Schematic of a mechanism that shows
633 why Pol α ChIP-ssSeq has a positive strand bias pattern. Red line: Watson strand, Green line:
634 Crick strand.

635

636 **Figure 4. Pol δ binds dsDNA at HU-stalled replication forks and preferentially binds**
637 **lagging-strand template ssDNA of active forks. (A)** A snap shot of Pol δ ChIP-ssSeq at three
638 origins (*ARS1511*, *ARS1512* and *ARS1513*) with HU or three time points without HU. G1-
639 synchronized yeast cells were released into fresh medium at 16°C without HU. Samples at the
640 indicated time points (72 min, 84min and 96min) samples were collected for Pol δ ChIP-ssSeq of
641 active forks. Sample was collected at 45 minutes after G1 release into HU at 30 °C for Pol δ
642 ChIP-ssSeq at HU-stalled forks. Please note that Pol δ ChIP efficiency was relatively low

643 compared to Pol ϵ and Pol α despite repeated attempts. **(B)** Pol δ ChIP-ssSeq peaks show positive
644 bias at active forks and no bias at HU-stalled forks. The average bias pattern of ChIP-ssSeq
645 peaks at 134 early replication origins is shown. The bias pattern indicated that Pol δ associates
646 with dsDNA, but associates with ssDNA more frequently at active forks **(C)** Dot-and-box plot
647 shows the bias pattern of Pol δ ChIP-ssSeq peaks at 134 individual early replication origins of
648 HU-stalled and active forks. Each dot represents one origin. **(D-E)** Schematics show the Pol δ -
649 DNA interaction at active forks **(D)** and HU-stalled forks **(E)**.

650

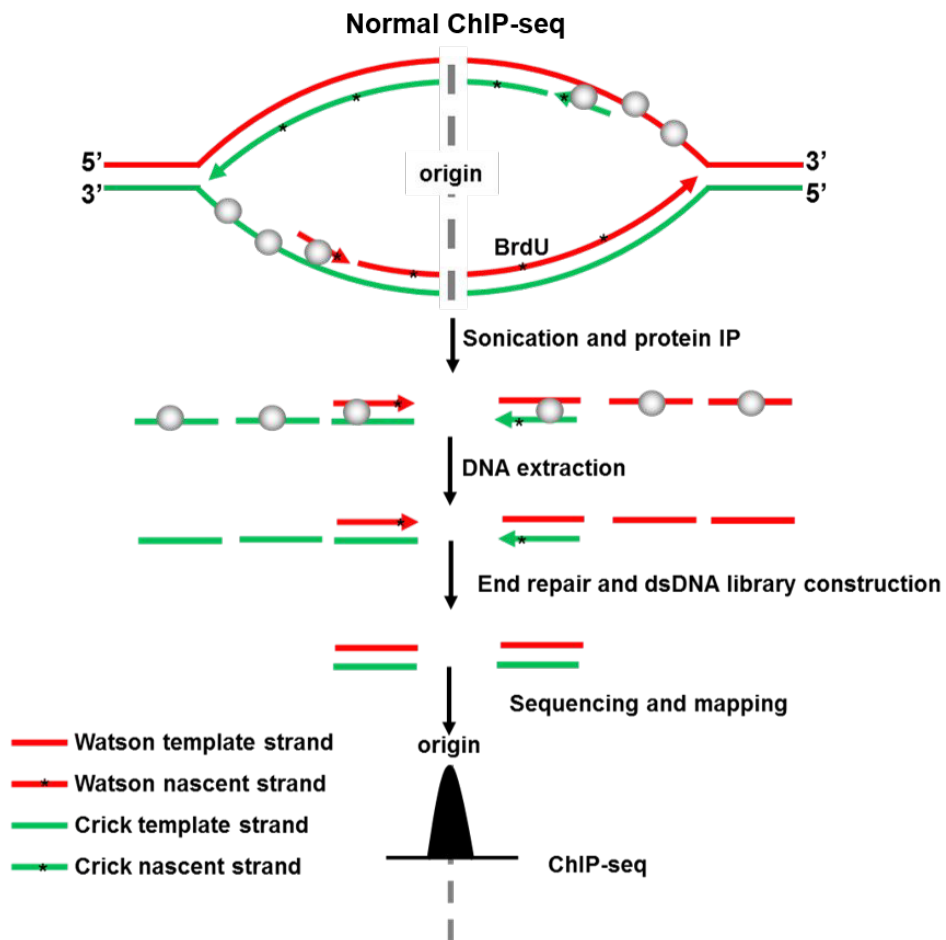
651 **Figure 5. Pol ϵ binds dsDNA at HU-stalled replication forks and preferentially binds**
652 **leading nascent ssDNA of active forks. (A-C)** Pol ϵ ChIP-ssSeq peaks show a positive bias
653 pattern at active forks and no bias at HU-stalled forks. The experiments were performed as
654 described in Fig. 4. except that Pol ϵ ChIP-ssSeq was performed. **(A)** a snap shot of Pol ϵ ChIP-
655 ssSeq peaks at *ARS1623*. **(B)** the average bias pattern of Pol ϵ ChIP-ssSeq peaks using a 200-bp
656 sliding window(early replication origins only). **(C)** Dot-and-box plot shows the bias pattern of
657 Pol ϵ ChIP-ssSeq peaks at 134 individual early origins. **(D-G)** The bias patterns of Pol ϵ ChIP-
658 ssSeq for HU-stalled and active fork are different. The bias patterns indicate that Pol ϵ associate
659 with dsDNA, but associates with ssDNA more frequently at active forks. **(D)** Flowchart of the
660 experimental procedure. Yeast cells arrested in G1 were released into HU for 45 minutes. A
661 fraction of cells were collected for Pol ϵ ChIP-ssSeq. The remaining cells were released into
662 fresh media after removal of HU. Samples were collected at the indicated time points after
663 release for Pol ϵ ChIP-ssSeq. **(E)** A snap shot of Pol ϵ ChIP-ssSeq peaks at *ARS1623* obtained
664 using HU arrested or released cells. **(F)** Analysis of the bias pattern of Pol ϵ ChIP-ssSeq peaks at

665 HU-stalled and active forks (3 time points after release). **(G)** Analysis of Pol ϵ ChIP-ssSeq peaks
666 at individual origins at active and HU-stalled forks. **(H-I)**. Schematics showing Pol ϵ at active **(H)**
667 and HU-stalled forks **(I)**. We propose that Pol ϵ binds to both template and nascent DNA of
668 active forks, but with a higher frequency to nascent DNA than template DNA, which leads to the
669 generation of Pol ϵ ChIP-ssSeq peak bias at active forks.

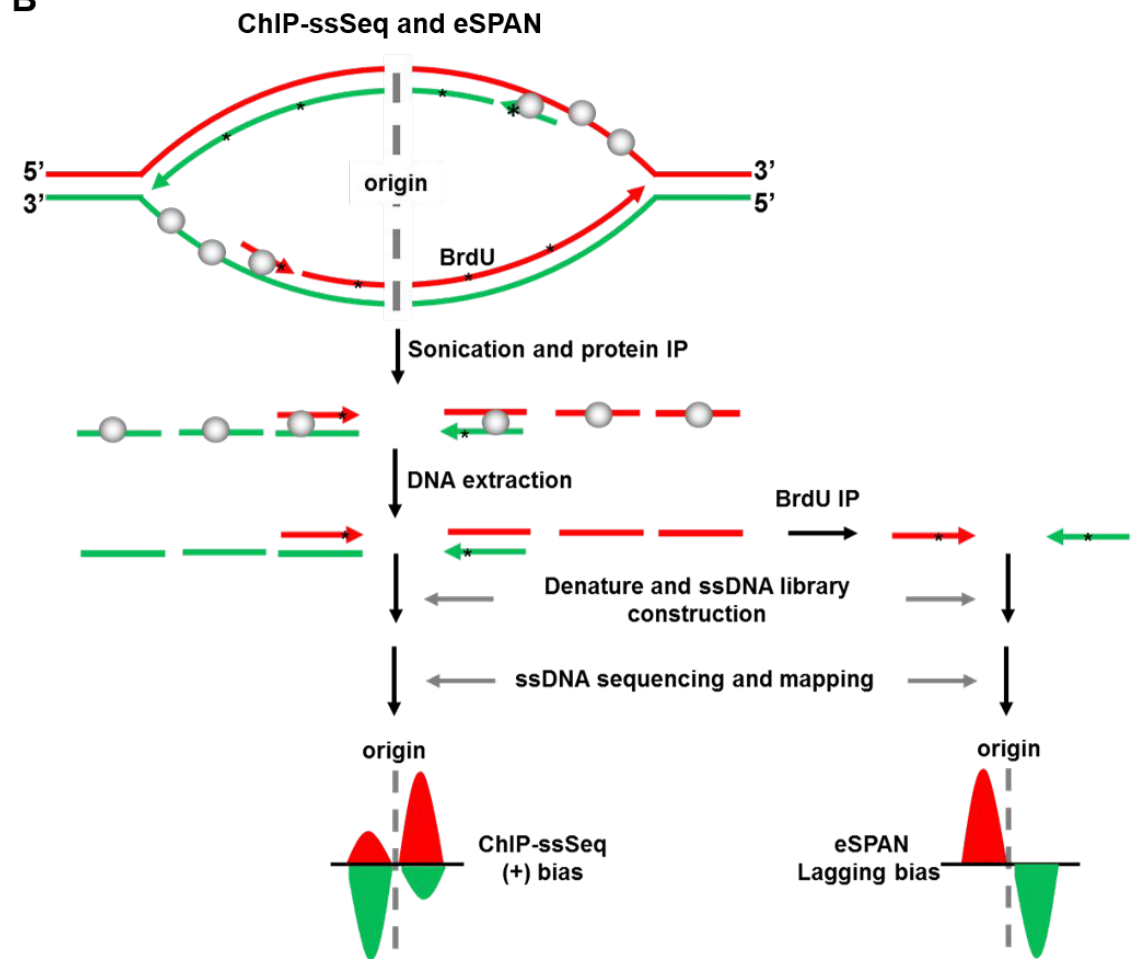
670

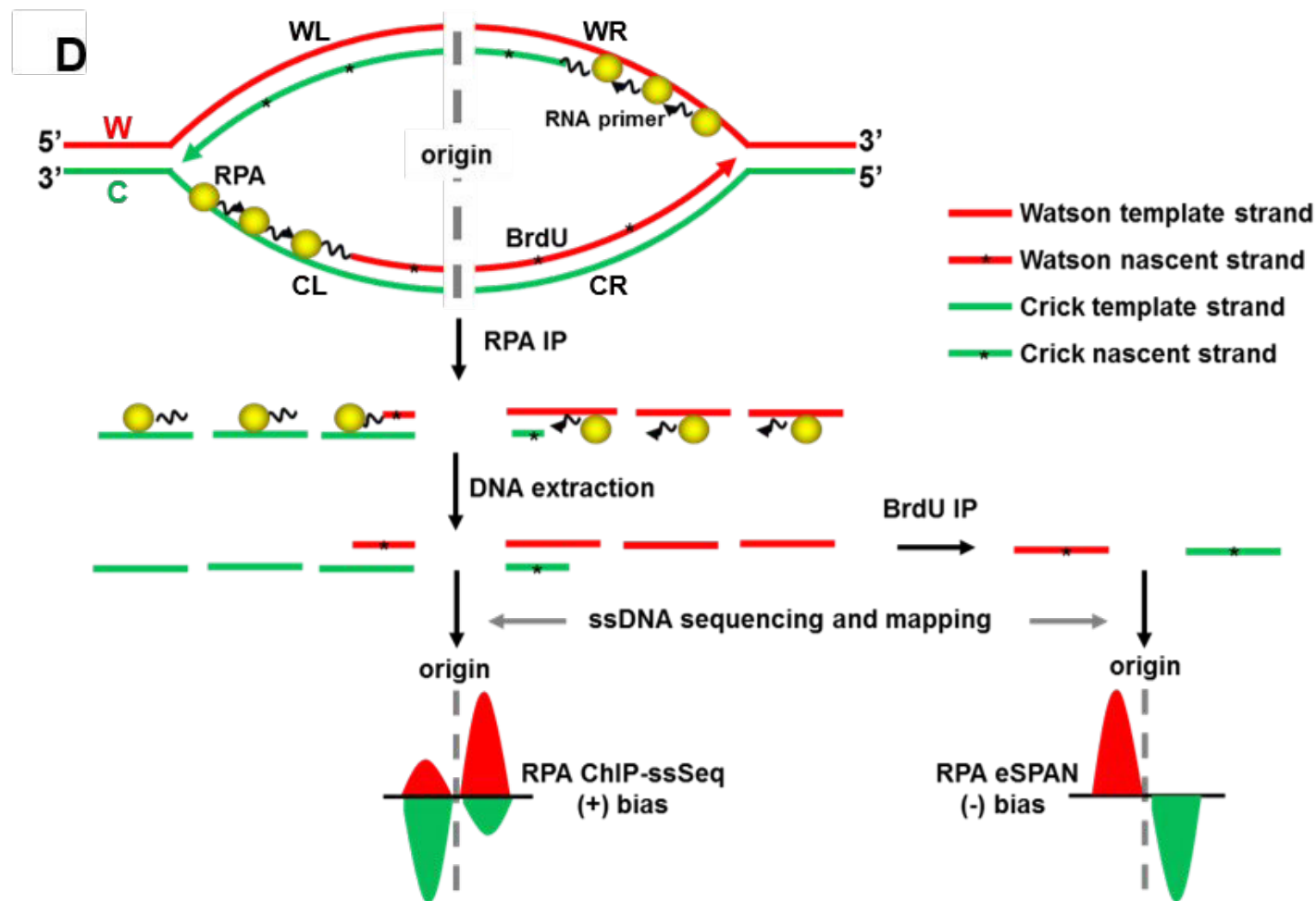
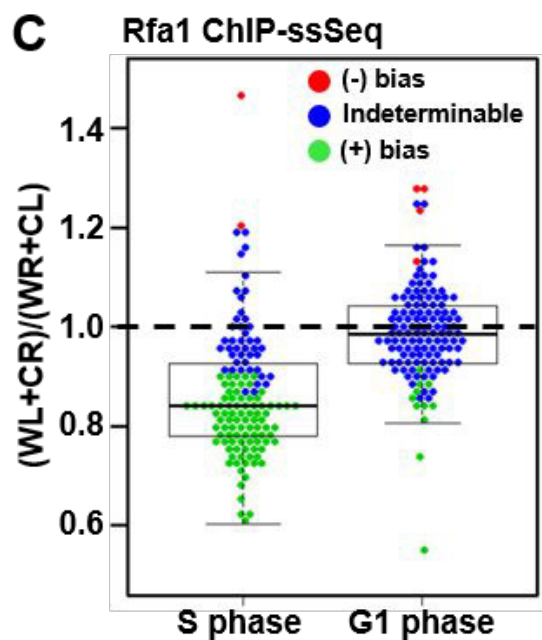
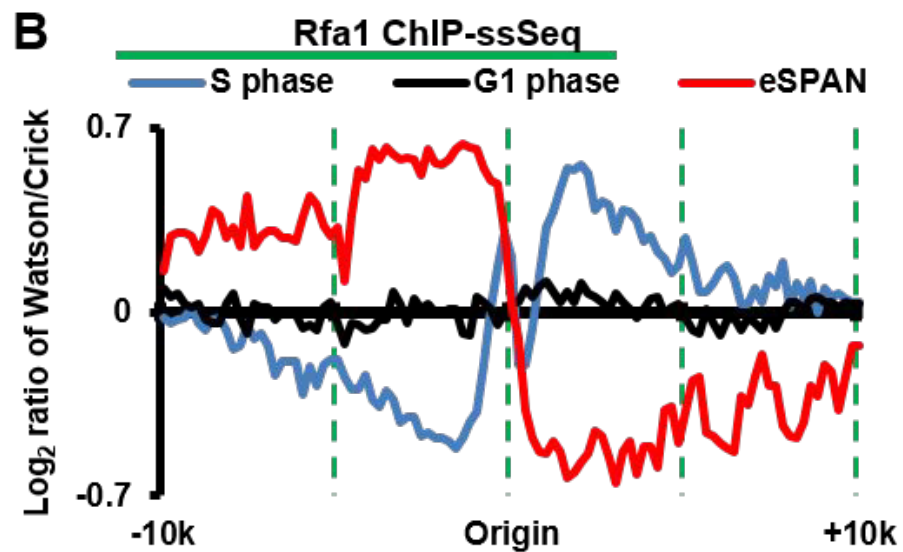
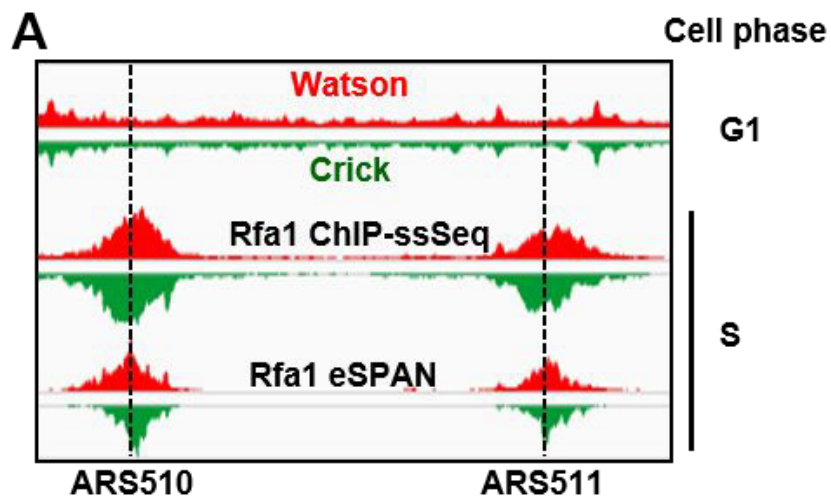
671

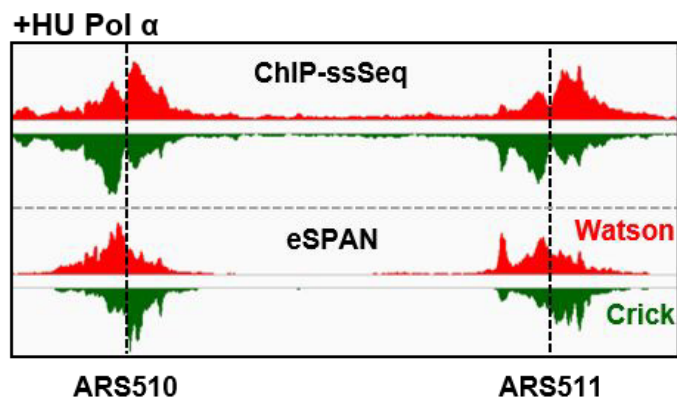
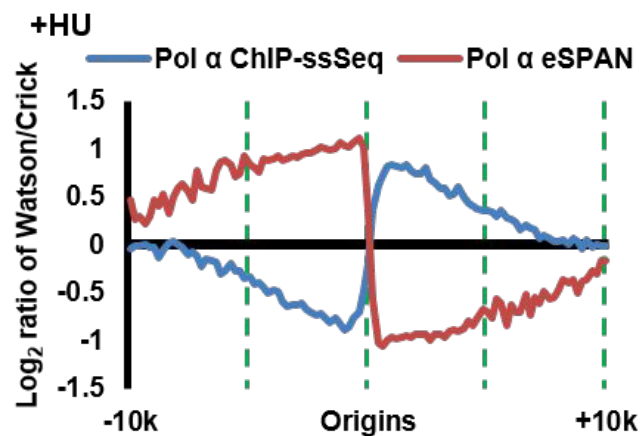
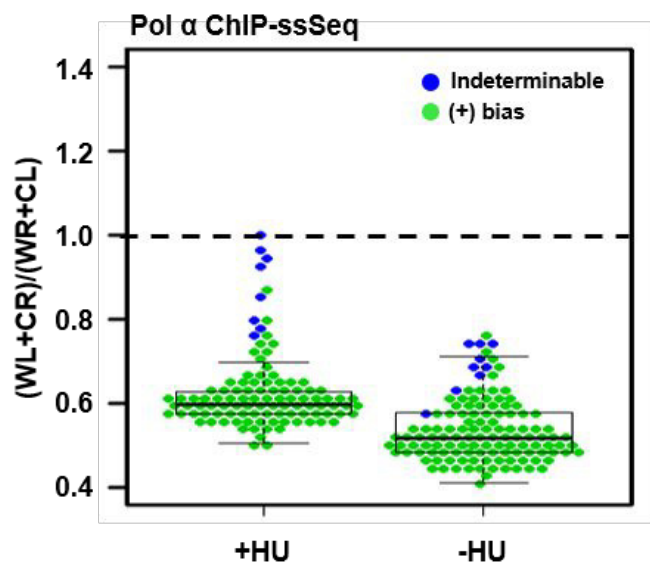
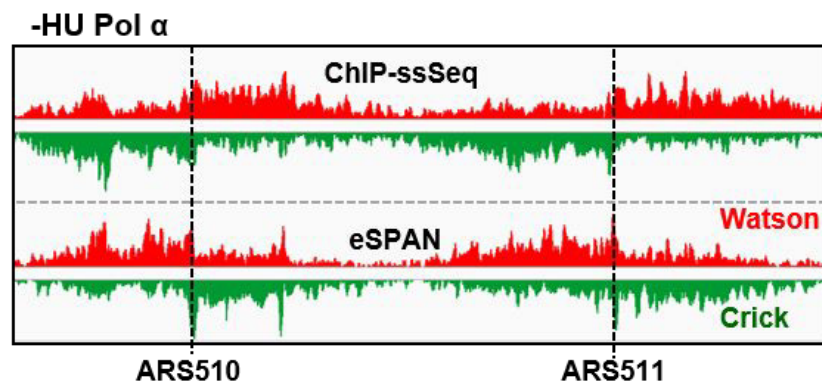
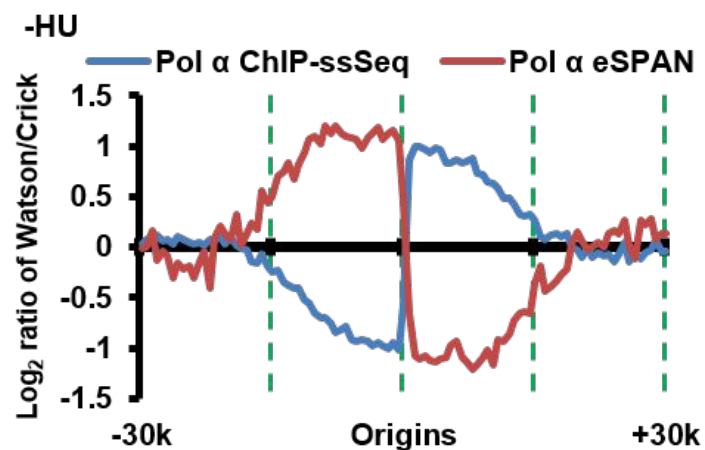
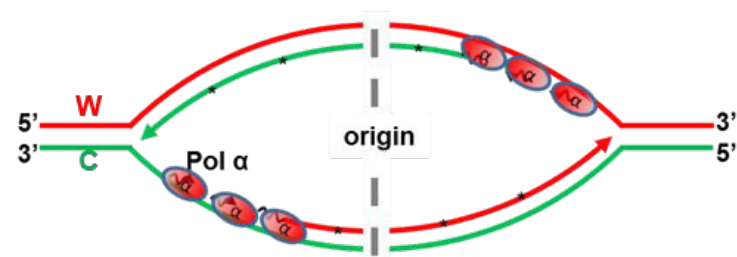
A

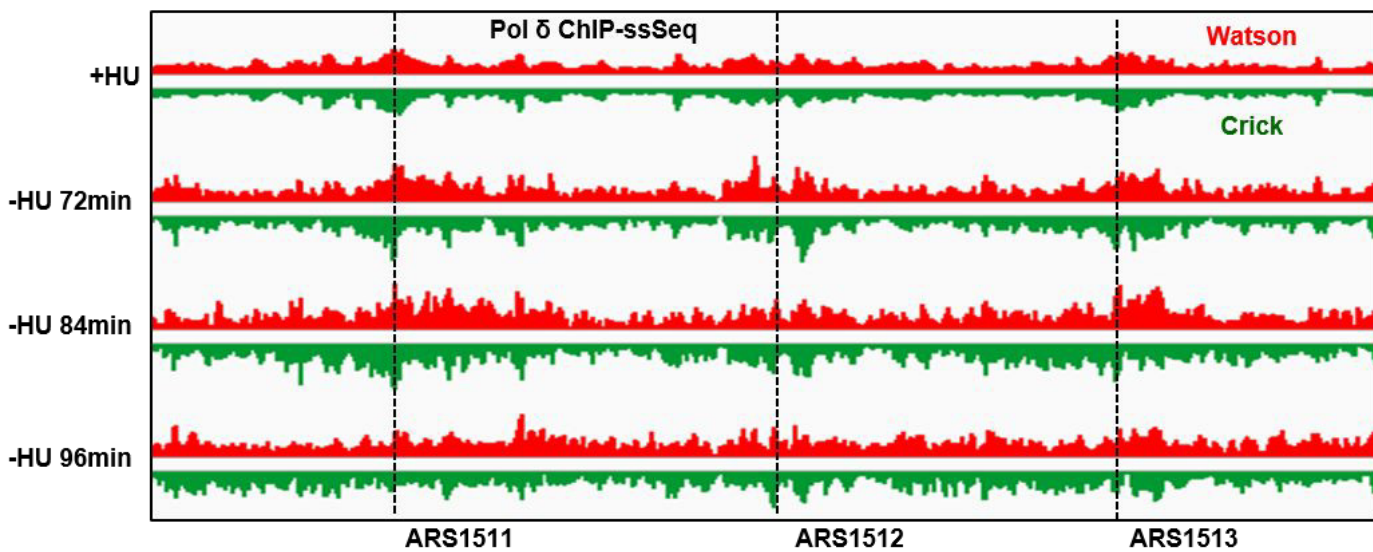
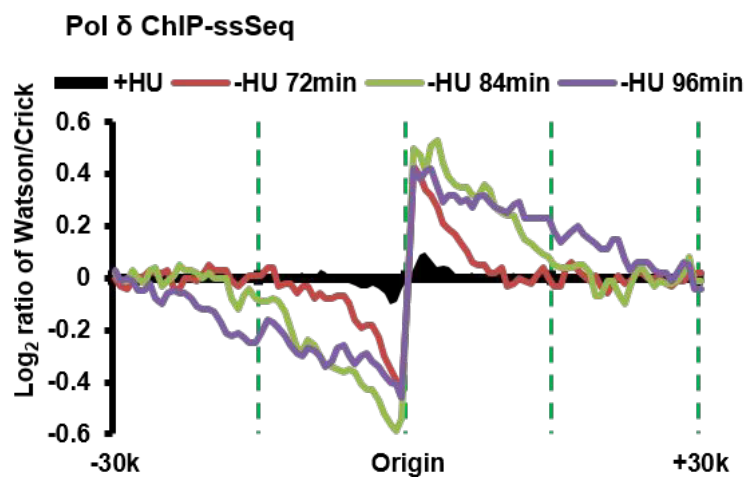
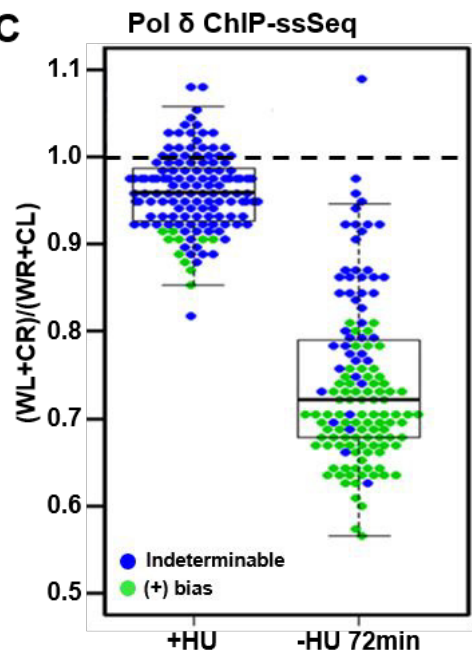


B

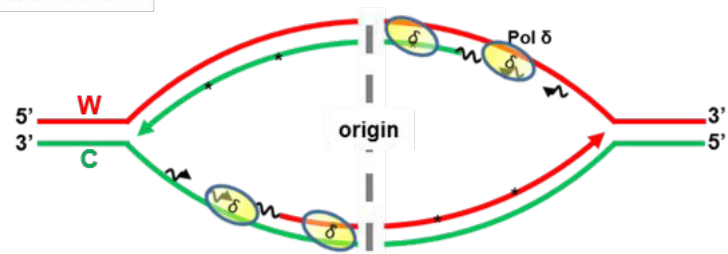




A**B****C****D****E****F**

A**B****C****D**

Normal fork

**E**

HU-stalled fork

